

T S4/5/1

4/5/1

DIALOG(R)File 351:Derwent WPI

(c) 2004 Thomson Derwent. All rts. reserv.

011228480 **Image available**

WPI Acc No: 1997-206383/199719

XRPX Acc No: N97-170327

**Synthetic unit storage method for speech rule synthesis - involves
storing information that relates syllable of each synthetic unit and
accent opposed position setting information in information file**

Patent Assignee: NTT DATA TSUSHIN KK (NITE)

Number of Countries: 001 Number of Patents: 001

Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
JP 8263090	A	19961011	JP 9560963	A	19950320	199719 B

Priority Applications (No Type Date): JP 9560963 A 19950320

Patent Details:

Patent No Kind Lan Pg Main IPC Filing Notes

JP 8263090 A 6 G10L-003/00

Abstract (Basic): JP 8263090 A

The method involves storing the information that relates the syllable of each synthetic unit and the accent opposed position setting information in an information file (25). A pair of detectors (21,22) are the accent opposed position setting information during search of the synthetic unit and select the desired synthetic unit.

The sound information is stored in a small storage unit. The silent section information are stored in a large storage unit. A recognition code of the small and large storage units are stored in a second information file (26).

ADVANTAGE - Enables efficient storage of local peak position information. Improves synthetic unit searching efficiency. Enables quick extraction of local peak position information.

Dwg.2/6

Title Terms: SYNTHETIC; UNIT; STORAGE; METHOD; SPEECH; RULE; SYNTHESIS;
STORAGE; INFORMATION; RELATED; SYLLABIC; SYNTHETIC; UNIT; ACCENTUATE;
OPPOSED; POSITION; SET; INFORMATION; INFORMATION; FILE

Index Terms/Additional Words: WAVEFORM; VOCAL; RULE; SYNTHESIS

Derwent Class: P86; W04

International Patent Class (Main): G10L-003/00

International Patent Class (Additional): G10L-005/04

File Segment: EPI; EngPI

?

This Page Blank (uspto)

(19)



JAPANESE PATENT OFFICE

PATENT ABSTRACTS OF JAPAN

JP4 8-263090

(11) Publication number: 08263090 A

(43) Date of publication of application: 11.10.96

(51) Int. Cl. G10L 3/00
G10L 5/04

(21) Application number: 07060963

(22) Date of filing: 20.03.95

(71) Applicant: N T T DATA TSUSHIN KK

(72) Inventor: KOYAMA TAKAO
MURAKAMI NORIYA
YOSHITANI AYANORI

(54) SYNTHESIS UNIT ACCUMULATING METHOD
AND SYNTHESIS UNIT DICTIONARY DEVICE

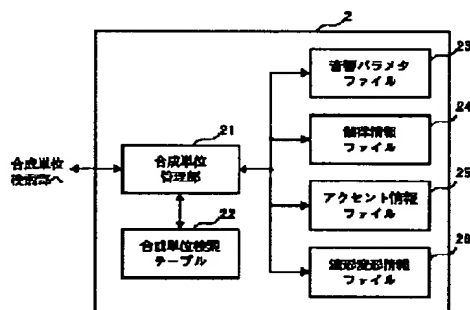
(57) Abstract:

PURPOSE: To provide a synthesis unit dictionary device in which the retrieval efficiency is increased during the selection of a synthesis unit that is to be used in a voice synthesis and local peak position information is efficiently accumulated.

CONSTITUTION: A synthesis unit dictionary device 2 consists of an accent information file 25 in which accent relative position information, that expresses the separation in terms of relative number of moras between the starting sound syllable of an individual synthesis unit and the accent position of a sentence paragraph that becomes the basis of the synthesis unit, is accumulated corresponding to the synthesis unit, synthesis unit selecting means 21 and 22 which select the candidates for the synthesis units employing the accent relative position information during the retrieving of the synthesis units and a waveform change information file 26 in which voiced sound interval information is accumulated in a relatively small accumulating unit among the voiced sound interval information defined by the interval of a local peak and voiceless sound-silence interval information, voiceless sound-silence interval information is accumulated in a

relatively large accumulating unit and an identification code of the accumulating unit is added to the prescribed part of each accumulating unit.

COPYRIGHT: (C)1996,JPO



This Page Blank (uspto)

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-263090

(43) 公開日 平成8年(1996)10月11日

(51) Int. Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
G10L 3/00			G10L 3/00	H
5/04			5/04	F
				E

審査請求 未請求 請求項の数5 O L (全6頁)

(21) 出願番号 特願平7-60963

(22) 出願日 平成7年(1995)3月20日

(71) 出願人 000102728

エヌ・ティ・ティ・データ通信株式会社
東京都江東区豊洲三丁目3番3号

(72) 発明者 小山 貴夫

東京都江東区豊洲三丁目3番3号 エヌ・
ティ・ティ・データ通信株式会社内

(72) 発明者 村上 憲也

東京都江東区豊洲三丁目3番3号 エヌ・
ティ・ティ・データ通信株式会社内

(72) 発明者 吉谷 文徳

東京都江東区豊洲三丁目3番3号 エヌ・
ティ・ティ・データ通信株式会社内

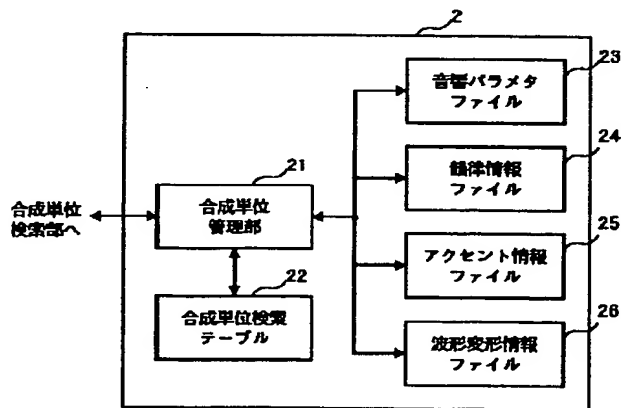
(74) 代理人 弁理士 鈴木 正剛

(54) 【発明の名称】 合成単位蓄積方法および合成単位辞書装置

(57) 【要約】

【目的】 音声合成で用いる合成単位を選定するときの検索効率を高め、かつローカルピーク位置情報を効率的に蓄積して合成単位辞書装置を提供する。

【構成】 個々の合成単位の開始音節が当該合成単位の基礎となる文節のアクセント位置から相対的に何モーラ離れているかを表すアクセント相対位置情報を当該合成単位に対応付けて蓄積したアクセント情報ファイル25と、合成単位の検索時にアクセント相対位置情報を用いて合成単位の候補を選定する合成単位選定手段21、22と、ローカルピークの間隔により定まる有声音区間情報と無声音・無音区間情報のうち相対的に小さい蓄積単位に有声音区間情報を蓄積するとともに相対的に大きい蓄積単位に無声音・無音区間情報を蓄積し、且つ各蓄積単位の所定部位に当該蓄積単位の識別符号を付して成る波形変形情報ファイル26を備えて合成単位辞書装置2を構成する。



【特許請求の範囲】

【請求項 1】 音声合成対象となる複数の合成単位を辞書装置に蓄積する際に、

個々の合成単位の開始音節が当該合成単位の基礎となる文節のアクセント位置から相対的に何モーラ離れているかを表すアクセント相対位置情報を当該合成単位に対応付けて蓄積することを特徴とする合成単位蓄積方法。

【請求項 2】 音声合成対象となる合成単位を辞書装置に蓄積する際に、前記合成単位の韻律制御に用いるローカルピーク位置情報を併せて蓄積する方法において、ローカルピークの間隔により定まる有声音区間情報と無声音・無音区間情報とにそれぞれ異なる蓄積単位を割り当て、かつ割り当てられた各蓄積単位の所定部位に当該蓄積単位の識別符号を付することを特徴とする合成単位蓄積方法。

【請求項 3】 前記有声音区間情報及び無声音・無音区間情報は、音声波形から切り出した合成単位を PCM 化して得た PCM データのローカルピーク間のサンプル数であることを特徴とする請求項 2 記載の合成単位蓄積方法。

【請求項 4】 合成対象となる複数の合成単位を蓄積して成る合成単位辞書装置において、個々の合成単位の開始音節が当該合成単位の基礎となる文節のアクセント位置から相対的に何モーラ離れているかを表すアクセント相対位置情報を当該合成単位に対応付けて蓄積したアクセント情報ファイルと、合成単位の検索時に前記アクセント情報ファイルに蓄積されたアクセント相対位置情報を用いて合成単位の候補を選定する合成単位選定手段と、を有することを特徴とする合成単位辞書装置。

【請求項 5】 合成対象となる複数の合成単位を蓄積して成る合成単位辞書装置において、ローカルピークの間隔により定まる有声音区間情報と無声音・無音区間情報のうち相対的に小さい蓄積単位に有声音区間情報を蓄積するとともに相対的に大きい蓄積単位に無声音・無音区間情報を蓄積し、且つ各蓄積単位の所定部位に当該蓄積単位の識別符号を付して成るローカルピーク位置情報ファイルを備えたことを特徴とする合成単位辞書装置。

【発明の詳細な説明】

【 0 0 0 1 】

【産業上の利用分野】 本発明は、規則音声合成に用いる音声合成単位辞書装置に関する。

【 0 0 0 2 】

【従来の技術】 従来の規則音声合成に用いる合成単位辞書装置は、音声波形から切り出した複数の合成単位と共に、各合成単位のピッチ周波数、経過時間、区間パワ等の韻律情報と、各合成単位の前後の音韻環境情報とを蓄積している。音声を作成する際には、上記韻律情報及び音韻環境をキー情報として所望の合成単位候補を索出し、索出した合成単位候補から目標とする韻律パタンとのパラメタ差を所定の評価式で評価して特定の合成単位

を決定している。

【 0 0 0 3 】 また、より自然な合成音声を生成する場合は、合成単位の接続位置において韻律パタンの不整合が起これないように、接続対象となる各合成単位の韻律を変形している。この韻律を変形する方法としては、例えば、広川、箱田、「波形編集型規則合成法におけるピッチ制御法の検討」平成 2 年 3 月音響学会講演論文集 1 - 4 - 7 に記載された「音声波形のローカルピークを用いたピッチ同期波形重畳法」等が知られている。

【 0 0 0 4 】

【発明が解決しようとする課題】 しかしながら、従来の合成単位辞書装置では、合成単位の蓄積量の増大に伴って合成単位の検索対象範囲が広がる。また、韻律パタンを評価基準とした場合には、数値的に似かよった合成単位が多くなり、優劣の判断が困難になる。その対策として前述の各パラメタに加えて前後数音韻の音韻環境情報を評価基準とすることも考えられるが、そうすると、音声合成単位辞書装置の検索効率が悪くなるという問題が生じる。

【 0 0 0 5 】 また、合成音声の自然性を確保するために前述の「ローカルピークを用いたピッチ同期波形重畳法」を実時間で実現する場合には、予め音声波形中からローカルピーク位置情報を抽出し、これを蓄えておく必要がある。しかし、合成単位の蓄積量が増加すると、それに比例して蓄積しておくべきローカルピーク位置情報も多くなる問題があった。

【 0 0 0 6 】 本発明の課題は、かかる問題点に鑑み、自然音声の韻律パタンに最も近い合成単位を選定するときの検索効率を高め、かつローカルピーク位置情報を少ない容量で効率的に蓄積する方法、及びこの方法により作成される合成単位辞書装置を提供することにある。

【 0 0 0 7 】

【課題を解決するための手段】 本発明は、まず、合成単位蓄積方法を提供する。この方法は、音声合成対象となる複数の合成単位を辞書装置に蓄積する際に、個々の合成単位の開始音節が当該合成単位の基礎となる文節のアクセント位置から相対的に何モーラ（音節の時間的まとまりの単位、以下同じ）離れているかを表すアクセント相対位置情報を当該合成単位に対応付けて蓄積することを特徴とする。

【 0 0 0 8 】 また、前記合成単位の韻律制御に用いるローカルピーク位置情報を併せて蓄積する際に、ローカルピーク（ピッチ構造における振幅の極大点、以下同じ）の間隔により定まる有声音区間情報と無声音・無音区間情報とにそれぞれ異なる蓄積単位を割り当て、かつ割り当てられた各蓄積単位の所定部位に当該蓄積単位の識別符号を付することを特徴とする。なお、有声音区間情報及び無声音・無音区間情報は、例えば音声波形から切り出した合成単位を PCM 化して得た PCM データのローカルピーク間のサンプル数の形で蓄積する。

【0009】本発明は、また、上記方法の実施により得られる合成単位辞書装置をも提供する。この装置は、合成対象となる複数の合成単位を蓄積して成る合成単位辞書装置において、個々の合成単位の開始音節が当該合成単位の基礎となる文節のアクセント位置から相対的に何モラ離れているかを表すアクセント相対位置情報を当該合成単位に対応付けて蓄積したアクセント情報ファイルと、合成単位の検索時に前記アクセント情報ファイルに蓄積されたアクセント相対位置情報を用いて合成単位の候補を選定する合成単位選定手段と、を有するものである。

【0010】本発明の他の構成に係る合成単位辞書装置は、合成対象となる複数の合成単位を蓄積して成る合成単位辞書装置において、ローカルピークの間隔により定まる有声音区間情報と無声音・無音区間情報のうち相対的に小さい蓄積単位に有声音区間情報を蓄積するとともに相対的に大きい蓄積単位に無声音・無音区間情報を蓄積し、且つ各蓄積単位の所定部位に当該蓄積単位の識別符号を付して成るローカルピーク位置情報ファイルを備えたことを特徴としている。

【0011】

【作用】一般に、文章はいくつかの発話単位に分割され、各発話単位は、一定の傾きで周波数の降下する特性（話調成分）の上に、文節固有のピッチ周波数、すなわちアクセント成分が重畳するモデルで表現される。個々のモデルは、各モラにおける音の高低の配置（ピッチパターン）によって記述されるが、日本語モデルのピッチパターンは、アクセント位置を頂点として略へ字状になる傾向がある。したがって、アクセント型の種類、つまり合成単位に対するアクセントの位置がわかれば、その合成単位の基礎となった文節のピッチパターンの概形を知ることが可能になる。また、文節のモラ数と音韻の継続時間長との間には相関があり、文節内のモラ数が多くなれば音韻の継続時間長が短くなる傾向にある。すなわち、モラ数とアクセント相対位置は、合成単位の選択の際の基準となるピッチ及び継続時間長をおおまかに反映しているため、これらの情報を有するアクセント相対位置情報を用いることにより、合成単位の継続時間長及びピッチ形状を大まかに加味した検索が可能になる。この場合、必要以上に長い前後の音韻環境を監視する必要が無い場合、合成単位を選択する際の探索区間を狭めることが可能となり、検索効率が高まる。

【0012】また、ローカルピーク位置情報を蓄積する際に、有声音区間情報を相対的に小さい蓄積単位で蓄積し、無声音・無音区間情報を大きな蓄積単位で蓄積することにより、一律の蓄積単位で蓄積する場合に比べて必要とする辞書容量が低減する。また、ローカルピークの絶対的な位置は、蓄積単位の識別符号から導かれる各区間情報の大きさ（データ長）を加算することで得られる。したがってローカルピークの位置や間隔を短い時間

で容易に導出することができる。

【0013】

【実施例】以下、図面を参照して本発明の合成音声単位辞書装置の好適な実施例を詳細に説明する。図1は、本発明の一実施例に係る規則音声合成装置の構成図である。まず、この規則音声合成装置1の全体的な処理概要を図1を参照して説明する。

【0014】規則音声合成装置1において、音声を作成する際には、合成しようとする文字列、アクセント型、及び韻律パターンを入力端子INに与える。入力端子INから入力された文字列等は、合成単位検索部3に送出される。合成単位検索部3では、入力された文字列等を、音声合成単位を検索する際に用いる所定のキー情報に変換し、該キー情報に基づいて合成単位辞書装置2から合成対象となる音声合成単位の候補素片と各候補素片の韻律情報及びアクセント相対位置情報とを索出する。合成単位検索部3では、また、索出した候補素片の韻律情報及びアクセント相対位置情報を適当な評価式を用いることにより1つに絞り込む。そして絞り込んだ素片の音響パラメタ及び波形変形情報を取得する。このようにして、合成に必要な合成単位を全て索出する。合成単位接続部4では、索出した合成単位を滑らかに接続してノイズの混入等を防止し、更に、韻律が目標の形状に近くなるように変形してこれを合成音声出力部5へ送出する。合成音声出力部5では、合成単位接続部4から送出された音響パラメタを音声信号に変換し、スピーカ等を介して出力する。

【0015】次に、本実施例による合成単位辞書装置2について説明する。この合成単位辞書装置2は、図2に示すように、上記合成単位検索部3へ双方向アクセス可能に接続された合成単位管理部21と、この合成単位管理部21へ双方向アクセス可能に接続された合成単位検索テーブル22、音響パラメタファイル23、韻律情報ファイル24、アクセント情報ファイル25、及び、波形変形情報ファイル26から構成される。

【0016】合成単位管理部21は、登録すべき情報を効率良く蓄積するとともに、他処理部に対し、要求に応じて情報を提供するためのインタフェースとして機能するものである。また、合成単位検索テーブル22は、所望の合成単位を高速に検索するためのテーブルである。音響パラメタファイル23は、合成音声を作成する際に用いる音響パラメタを格納する領域である。韻律情報ファイル24は、各合成単位の平均ピッチ周波数、継続時間長、区間パワー及び隣接音韻の種類等の情報を格納する領域である。なお、本実施例では、PCM (Pulse code modulation) 方式により音声波形をデジタル化したものを格納しているため、上記各種パラメタはこれらデジタルデータ (PCMデータ) に対応したものである。

【0017】アクセント情報ファイル25は、個々の合

成単位が何モーラの文節から切り出され、且つその合成単位が当該文節のアクセント位置から相対的に何モーラ離れているかを表すアクセント相対位置情報を合成単位別に格納するファイルである。

【0018】このアクセント情報ファイル25の詳細を図3及び図4を参照して説明する。前述のように、合成単位に対するアクセントの相対位置がわかれば、その合成単位の基礎となった文節のピッチパタンの概形を知ることが可能になる。例えばモーラ数が3の合成単位の場合は、図3(a)～(d)に示すように、平板型の3モーラ0型、先頭にアクセントがある3モーラ1型、中央部分にアクセントがある3モーラ2型、及び末尾にアクセントがある3モーラ3型では、各合成単位の先頭位置に対するアクセント相対位置がわかれば、当該合成単位の基礎となった文節のピッチパターンがわかる。そこで、本実施例では、先頭位置がアクセント位置である場合はアクセント相対位置を'0'、同様に、一音節前であれば' -1'、一音節後ろであれば' +1'として定める。また、アクセントのない平板型では便宜的に2モーラ目をアクセント位置として定める。そしてこの情報を合成単位毎に付加する。このようにして作成したアクセント情報ファイル25のテーブルイメージ例401を図4に示す。

【0019】図4において、合成単位番号は、本実施例の規則音声合成装置1が各合成単位に対して単一に与える値であり、モーラ数及びアクセント相対位置を一意に対応付けている。したがって、合成単位番号を検索する際にモーラ数やアクセント相対位置を用いることができ、合成単位の選択の際の基準となるピッチ及び継続時間長をもおおまかに知ることが可能となる。

【0020】次に、波形変形情報ファイル26の詳細内容を図5及び図6を参照して説明する。有声音区間において、音声波形は周期的な形状をなしている。これは一般にピッチ構造と呼ばれる。各ピッチにおいては、図5に示すように、周期的に振幅の大きなピーク(502, 503, 504)が現われる。これらピークをローカルピークと称する。音声のピッチ周期は、ほぼ2.5ms～10msの間に分布しており、周期が短くなるほど声が高くなり、逆に周期が長くなるほど声が低く聞こえる。よって、もとの音声のピッチ周期を変更することで、声の高さを制御することが可能となる。実際の音声波形におけるピッチ間隔は図5下段に示すとおりであり、図示の例では、3600サンプル505、97サンプル506、及び92サンプル507の場合を示している。各ピッチ間隔は、サンプリング周波数10kHzの音声を用いた場合の例である。また、ピッチ単位で間引きや繰り返し処理を行うことで、合成単位の継続時間長の制御が可能となる。これらの韻律制御を行う手法が前述のピッチ同期波形重畳法であり、この手法では、ローカルピーク位置をピッチ毎の目印として用いている。

【0021】図6は、波形変形情報ファイル26におけるローカルピーク位置情報の格納要領、すなわちローカルピーク位置情報ファイルの説明図である。本実施例では、ローカルピーク位置情報は、PCMデータにおけるピーク間のサンプル数で表す。但し、先頭のローカルピーク位置は、音声データの始端とピーク位置との間隔である。このようにすれば、ローカルピークの絶対的な位置は、先頭からのピーク間隔を加算することで得られる。ピーク間隔の傾向は、大きく分けると有声音区間の短い区間のものと無声音または無音区間における長めの区間のものとがある。有声音区間では、ピッチ周期が2.5ms～10ms程度であり、サンプリング周波数を10kHzとしたとき、25～100sampleの間隔でローカルピークが出現する。一方、無声音区間は、一般的に100ms～1s程度であり、ローカルピークの出現間隔は、1000～10000sample程度である。

【0022】そこで、本実施例では、ローカルピーク位置情報ファイル(波形変形情報ファイル26)内に短い蓄積単位と長い蓄積単位とを用意し、各ピーク間隔に合わせて蓄積単位の長さを選択する。図6に示す例では、(a)に示す短い蓄積単位601に8ビットを割り付ける。但し、先頭の1ビットは蓄積単位長を表すフラグ603であり、論理0の時は短い蓄積単位を表し、残り7ビット602を割り付ける。ピーク間隔の表現能力としては、0～12sampleを表現することが可能である。また、ピーク間隔として余りにも短いもの(例えば0～2ms程度)はあり得ないとして、表現できる値に定数値を割り当てることで更に効率化が可能である。一方、(b)に示す長い蓄積単位604には16ビットを割り当てる。長い蓄積単位でも、先頭の1ビットを蓄積単位長を表すフラグ606として用いる。この場合は、論理1を長い蓄積単位の識別符号として用いる。そして、残り15ビット605を割り当てる。ピーク間隔の表現能力としては、0～32786sampleを表現することが可能である。

【0023】実際に使用する合成単位607においては、(c)に示すように、先ず先頭ビットのフラグ608を確認して取得するデータの長さを識別する。ここで、ピーク間隔を取得し、最初に読み込んだデータ長610から、次のデータの読み出し位置(次のデータの先頭フラグ)609を計算する。この処理を随時行うことで、一連のローカルピーク位置情報を読み出すことが可能となる。また、ローカルピーク位置は、合成単位毎に与えられており、各合成単位での終端位置は、一定値を検出したときの終了値とする。また、始端の読み出し位置(アドレス)は、所定の管理情報を用いて指定する。

【0024】このように、アクセント情報ファイル25を用いることで、ピッチ及び継続時間長を考慮した検索が可能となり、合成単位の候補の絞り込む際の探索区間

を大幅に狭めることが可能となる。アクセント情報ファイル 25 は、継続時間長に大きな影響を与えるモーラ数をも保持しているため、これを探索時の評価基準として用いることにより選択される合成単位の継続時間長が自然音声に近いものとなる。また、ローカルピーク位置情報を迅速に索出できるため、韻律の制御を実時間で行うことが可能であり、さらに、ローカルピークの間隔に応じて適当な蓄積単位を割り当てているので、単純に同一の蓄積単位を与える場合の 50%、25% 程度の容量で同一情報の蓄積が可能となる。これにより波形辞書のサイズを従来より格段に小さくすることが可能となる。

【0025】なお、本実施例で用いたモーラ数やアクセント相対位置、およびピッチ間隔のサンプル数は例示であり、必ずしも本実施例の数値に限定されるものではない。

【0026】

【発明の効果】以上の説明から明らかなように、本発明の合成単位蓄積方法及びこれにより得られた合成単位辞書装置によれば、波形規則音声合成装置で用いる大量の合成単位の中から、自然音声の韻律パターンに最も近い合成単位を選定する際に、おおまかな探索範囲を示唆することが可能になり、検索効率が格段に高まる効果がある。

【0027】また、音声波形のローカルピーク位置情報を効率良く蓄積し得る効果があり、しかも必要なローカルピーク位置情報を迅速に抽出することができるので、例えば合成音声の発声速度や声の高さを実時間で変更す

る用途において顕著な効果を奏することができる。

【図面の簡単な説明】

【図 1】本発明の一実施例に係る合成単位辞書装置を用いた音声合成装置の一例を示すブロック構成図。

【図 2】本実施例の合成単位辞書装置のブロック構成例を示す図。

【図 3】本実施例によるモーラとアクセント相対位置との関係説明図。

【図 4】本実施例によるアクセント情報ファイルのテーブルイメージ説明図。

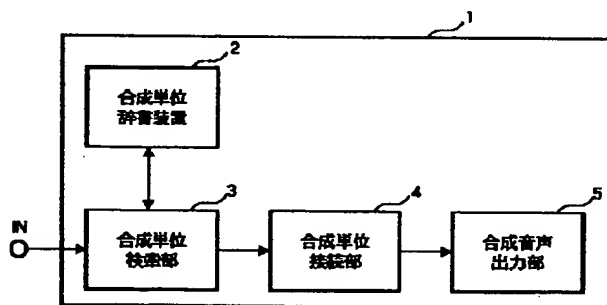
【図 5】ローカルピーク及びピーク間情報の説明図。

【図 6】本実施例による波形変形情報ファイルにおけるローカルピーク位置情報の蓄積状態説明図。

【符号の説明】

- 1 規則音声合成装置
- 2 合成単位辞書装置
- 3 合成単位検索部
- 4 合成単位接続部
- 5 合成音声出力部
- 20 21 合成単位管理部
- 22 合成単位検索テーブル
- 23 音響パラメタファイル
- 24 韻律情報ファイル
- 25 アクセント情報ファイル
- 26 波形変形情報ファイル

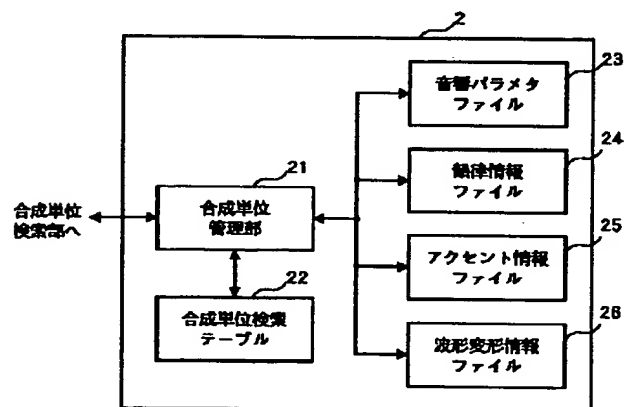
【図 1】



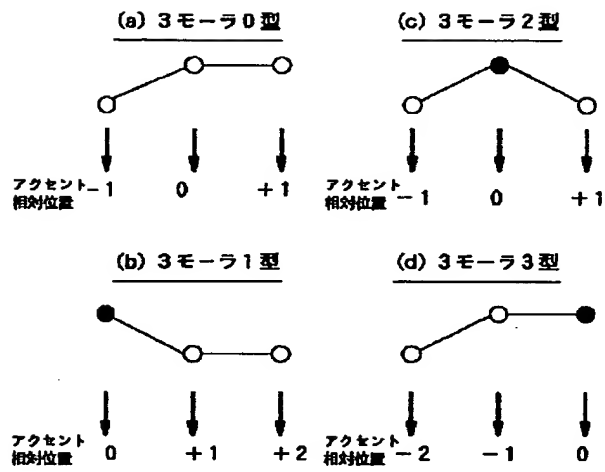
【図 4】

合成単位番号	モーラ数	アクセント相対位置
101	6	-1
102	4	0
103	8	2
⋮	⋮	⋮
⋮	⋮	⋮

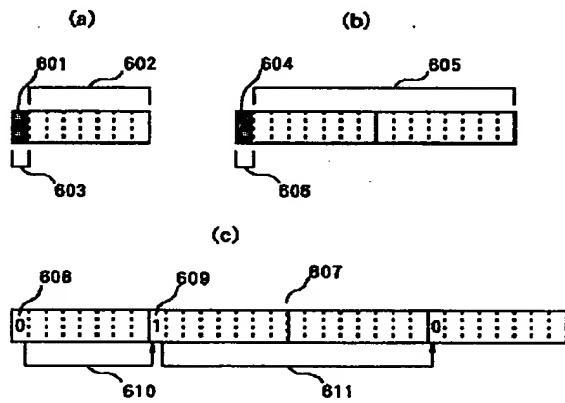
【図 2】



【図3】



【図6】



【図5】

